

DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ALPHA SUBUNIT OF BETA CONGLYCININ.
OS GLYCINE max (Soybean).
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta;
OC Magnoliophyta; eudicotyledons: core eudicots: Rosidae; eustoids 1;
OC Fabales; Fabaceae; Papilionoideae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. WASESUZUNARI;
RA Maryama N., Katsube T., Wada Y., Utsumi S.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008678; BAA23360.1; -
DR HSSP; P50477; 1CAU.
DR MENDEL; 25074; GLYma;1188;25074.
DR INTERPRO; IPR001113; -
DR PFAM; PF00546; Seedstore_7s; 1.
DR PRODOM; PD081059; -; 1.
SQ SEQUENCE 544 AA; 63296 MW; 048B243929A9C9AB CRC64;

Query Match 41.6%; Score 131; DB 10; Length 544;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 VEKECEGEIIPRRPQHPER 55
ID 096456 PRELIMINARY; PRT; 556 AA.
AC 096456;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE SPOROZOITE SURFACE PROTEIN 2.
GN SSP2.
OS Plasmodium vivax.
OC Eukaryota: Alveolata: Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SAL I;
RA Rogers W.O., Gowda K., Saccl J., Hoffman S.L.;
RT "Construction and immunogenicity of DNA vaccine plasmids encoding four
RL Plasmodium vivax candidate vaccine antigens.";
DR EMBL; AF063136; AAC97484.1; -
DR INTERPRO; IPR000884; -
DR INTERPRO; IPR002035; -
DR PFAM; PF00090; tsp-1; 1.
DR PFAM; PF00092; vwa-1.
DR PRINTS; PR00453; VWFADOMAIN.
FT VARIANT 134 D -> E.
FT VARIANT 160 V -> I.
FT VARIANT 166 R -> T.
DE 172 N -> K.
FT VARIANT 268 D -> G.
FT VARIANT 352 K -> E.
FT VARIANT 366 T -> A.
FT VARIANT 391 I -> R.
FT VARIANT 464 E -> Q.
FT VARIANT 494 T -> I.
SQ SEQUENCE 556 AA; 61208 MW; 530B5A0C86E8F3A7 CRC64;

Query Match 22.1%; Score 69.5; DB 5; Length 556;
Best Local Similarity 33.3%; Pred. No. 1.2;
Matches 18; Conservative 4; Mismatches 23; Indels 9; Gaps 2;

QY 8 CLOSNERDSYRNOACHARNLKVKECEGE-----IPRRPQHPER 54
ID 096456 PRELIMINARY; PRT; 556 AA.
AC 096456;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE SPOROZOITE SURFACE PROTEIN 2.
GN SSP2.
OS Plasmodium vivax.
OC Eukaryota: Alveolata: Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SAL I;
RA Rogers W.O., Gowda K., Hoffman S.L.;
RT "Construction and immunogenicity of DNA vaccine plasmids encoding four
RL Plasmodium vivax candidate vaccine antigens.";
DR EMBL; AF063137; AAC97485.1; -
DR INTERPRO; IPR000884; -
DR INTERPRO; IPR002035; -
DR PFAM; PF00090; tsp-1; 1.
DR PFAM; PF00092; vwa-1.
DR PRINTS; PR00453; VWFADOMAIN.
FT VARIANT 134 D -> E.
FT VARIANT 160 V -> I.
FT VARIANT 166 R -> T.
DE 172 N -> K.
FT VARIANT 268 D -> G.
FT VARIANT 352 D -> E.
FT VARIANT 391 M -> R.
FT VARIANT 464 E -> Q.
SQ SEQUENCE 556 AA; 61208 MW; 7FAFE294B3E32AA7 CRC64;

Query Match 22.1%; Score 69.5; DB 5; Length 556;
Best Local Similarity 33.3%; Pred. No. 1.2;
Matches 18; Conservative 4; Mismatches 23; Indels 9; Gaps 2;
QY 8 CLOSNERDSYRNOACHARNLKVKECEGE-----IPRRPQHPER 54
ID 096456 PRELIMINARY; PRT; 556 AA.
AC 096456;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE SPOROZOITE SURFACE PROTEIN 2.
GN SSP2.
OS Plasmodium vivax.
OC Eukaryota: Alveolata: Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SAL I;
RA Rogers W.O., Gowda K., Hoffman S.L.;
RT "Construction and immunogenicity of DNA vaccine plasmids encoding four
RL Plasmodium vivax candidate vaccine antigens.";
DR EMBL; AF063137; AAC97485.1; -
DR INTERPRO; IPR000884; -
DR INTERPRO; IPR002035; -
DR PFAM; PF00090; tsp-1; 1.
DR PFAM; PF00092; vwa-1.
DR PRINTS; PR00453; VWFADOMAIN.
FT VARIANT 134 D -> E.
FT VARIANT 160 V -> I.
FT VARIANT 166 R -> T.
DE 172 N -> K.
FT VARIANT 268 D -> G.
FT VARIANT 352 D -> E.
FT VARIANT 391 M -> R.
FT VARIANT 464 E -> Q.
SQ SEQUENCE 556 AA; 61208 MW; 7FAFE294B3E32AA7 CRC64;

Query Match 22.1%; Score 69.5; DB 5; Length 556;
Best Local Similarity 33.3%; Pred. No. 1.2;
Matches 18; Conservative 4; Mismatches 23; Indels 9; Gaps 2;
QY 8 CLOSNERDSYRNOACHARNLKVKECEGE-----IPRRPQHPER 54
ID 096456 PRELIMINARY; PRT; 556 AA.
AC 096456;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE SPOROZOITE SURFACE PROTEIN 2.
GN SSP2.
OS Plasmodium vivax.
OC Eukaryota: Alveolata: Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SAL I;
RA Rogers W.O., Gowda K., Hoffman S.L.;
RT "Construction and immunogenicity of DNA vaccine plasmids encoding four
RL Plasmodium vivax candidate vaccine antigens.";
DR EMBL; AF063137; AAC97485.1; -
DR INTERPRO; IPR000884; -
DR INTERPRO; IPR002035; -
DR PFAM; PF00090; tsp-1; 1.
DR PFAM; PF00092; vwa-1.
DR PRINTS; PR00453; VWFADOMAIN.
FT VARIANT 134 D -> E.
FT VARIANT 160 V -> I.
FT VARIANT 166 R -> T.
DE 172 N -> K.
FT VARIANT 268 D -> G.
FT VARIANT 352 D -> E.
FT VARIANT 391 M -> R.
FT VARIANT 464 E -> Q.
SQ SEQUENCE 556 AA; 61208 MW; 7FAFE294B3E32AA7 CRC64;

[illegible]

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003576; AAF51000.1; -.
 DR HSSP: P35535; 1EMN.
 DR FLYBASE: FBgn0031612; CG15637.
 DR INTERPRO: IPR000152; -.
 DR INTERPRO: IPR000561; -.
 DR INTERPRO: IPR001881; -.
 DR Pfam: PF00008; EGF_28.
 DR PROSITE: PS00010; ASX_HYDROXYL; 18.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 18.
 DR PROSITE: PS01187; EGF_CA; 17.
 SO SEQUENCE 3680 AA; 388637 MW; E5E972E1A3479EFF CRC64;

Query Match 21.1%; Score 66.5; DB 5; Length 3680;
 Best Local Similarity 27.4%; Pred. No. 15;
 Matches 17; Conservative 8; Mismatches 20; Indels 17; Gaps 3;

OY 5 HNCGLSCN-----SEDSYRNQAC-----HARCNLKVEKE-ECGEGELPRPR 47
 DB 2818 HNCVYGCIVHDDDCSASESCRNDKCVNPLENCGPNAACVSNIHRASCSLESMVNPPT 2877
 OY 48 PR 49
 DB 2878 PQ 2879

RESULT 8 PRELIMINARY: PRT; 963 AA.
 ID O9M8T1
 AC O9M8T1:
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE F13E7.16 PROTEIN.
 GN F13E7.16.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 RA Romling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F13E7 genomic sequence.";
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC018363; AAF26970.1; -.
 SO SEQUENCE 963 AA; 105391 MW; D204IDC304280339 CRC64;

Query Match 21.0%; Score 66; DB 10; Length 963;
 Best Local Similarity 32.3%; Pred. No. 5.3;
 Matches 21; Conservative 3; Mismatches 13; Indels 28; Gaps 3;

OY 1 ENPKHNKCL-----QSCNSERDSYRNQACHARCNLKVEKE-----EC 38
 DB 832 ENPNSNKSLLDRTLMSQSCMKEDAKESKACSG-----TEKENAFSVSYGEGVAVET 885
 OY 39 EGGEL 43
 DB 886 EGGEL 890

RESULT 9

O94727
 ID O94727 PRELIMINARY: PRT; 510 AA.
 AC O94727:
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE THROMBOSPONDIN-RELATED ANONYMOUS PROTEIN (FRAGMENT).
 GN TRAP.
 OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5855;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SAI-1;
 RX MEDLINE=97193932; PubMed=9041517;
 RA Templeton T.J., Kaslow D.C.;
 RT "Cloning and cross-species comparison of the thrombospondin-related
 RT anonymous protein (TRAP) gene from Plasmodium knowlesi, Plasmodium
 RT vivax and Plasmodium gallinaceum.";
 RL Mol. Biochem. Parasitol. 84:13-24(1997).
 DR EMBL: U64901; AAC47463.1; -.
 DR INTERPRO: IPR000884; -.
 DR INTERPRO: IPR002035; -.
 DR Pfam: PF00090; tSP_1; 1.
 DR Pfam: PF00092; vwa; 1.
 DR PRINTS: PR00453; VWFADOMAIN.
 FT NON_TER 1
 FT NON_TER 510
 SO SEQUENCE 510 AA; 55601 MW; 72687C7BAFDFAID CRC64;

Query Match 20.8%; Score 65.5; DB 5; Length 510;
 Best Local Similarity 33.3%; Pred. No. 3.5;
 Matches 18; Conservative 3; Mismatches 24; Indels 9; Gaps 2;

OY 8 CLOSCHSDSYRNQACHARCNLKVEKECEGE-----IPRRPRPQHPH 54
 DB 210 CSVTGGRGTHSRSPSLHRCCTTHMV--SECEGEGVPEPEPLVPAPLPYPE 261

RESULT 10
 ID O42918 PRELIMINARY: PRT; 774 AA.
 AC O42918:
 DT 01-JAN-1999 (Tremblrel. 09, Created)
 DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE PUTATIVE ALPHA-AMYLASE C16A3.13 PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-
 DE GLUCAN GLUCANOHYDROLASE).
 GN SPBCL6A3.13.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
 OC Schizosaccharomycetaceae; Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wood V., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.;
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: TO THE FAMILY 13 OF GLUCOSYL HYDROLASES, ALSO KNOWN AS
 CC THE ALPHA-AMYLASE FAMILY.
 DR EMBL: A1021748; CAI16864.1; -.
 DR HSSP: P10529; 7TAA.
 DR INTERPRO: IPR000461; -.
 DR Pfam: PF00128; alpha-amylase; 1.
 KW Hypothetical protein; Hydrolase; Glycosidase; Carbohydrate metabolism;
 KW Calcium; Glycoprotein; Signal.
 FT SIGNAL 1
 FT CHAIN 25 774 PUTATIVE ALPHA-AMYLASE C16A3.13.

```

FT ACT_SITE 461 461 BY SIMILARITY.
FT ACT_SITE 553 553 BY SIMILARITY.
FT DISULFID 52 60 BY SIMILARITY.
FT DISULFID 229 314 BY SIMILARITY.
FT DISULFID 495 539 BY SIMILARITY.
FT DISULFID 690 724 BY SIMILARITY.
FT CARBOHYD 143 143 POTENTIAL.
FT CARBOHYD 187 187 POTENTIAL.
FT CARBOHYD 214 214 POTENTIAL.
FT CARBOHYD 243 243 POTENTIAL.
FT CARBOHYD 272 272 POTENTIAL.
FT CARBOHYD 299 299 POTENTIAL.
FT CARBOHYD 326 326 POTENTIAL.
FT CARBOHYD 355 355 POTENTIAL.
FT CARBOHYD 382 382 POTENTIAL.
FT CARBOHYD 409 409 POTENTIAL.
SQ SEQUENCE 774 AA: 89459 MW: 92CEAAE45664269B CRC64;

```

```

Query Match 20.8%; Score 65.5; DB 3; Length 774;
Best Local Similarity 28.1%; Pred. No. 5;
Matches 16; Conservative 8; Mismatches 20; Indels 13; Gaps 3;

```

```

OY 6 NKLQSCNSRSDYRNQACHRCMLKVE-----KECEEGELPRPRP----RPOH 52
Db 167 HKCEHSCSHDKFPRPVH--NGTRDHRPKWKEHCHHGKPRPVPHNGTRKPDH 220

```

RESULT 11

```

ID 008641 PRELIMINARY; PRT; 340 AA.
AC 008641;

```

```

DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE RAY PROTEIN.
GN SH3YL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL6; TISSUE=SKIN;
RA Aoki N.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: D85926; BAA19679.1; -.
DR HSSP: P29354; 1GFC.
DR MCD: MGI:1346118; Sh3Y11.
DR INTERPRO: IPR001452; -.
DR PFAM: PF00018; SH3; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR PROSITE: PS50002; SH3; 1.
SQ SEQUENCE 340 AA: 37028 MW: 3FE12C0A8530C74E CRC64;

```

```

Query Match 20.5%; Score 64.5; DB 11; Length 340;
Best Local Similarity 38.6%; Pred. No. 3.2;
Matches 17; Conservative 8; Mismatches 14; Indels 5; Gaps 3;

```

```

OY 13 NSERDSTRNQACHRCMLKVEKECEEGELPRP--RPPPOH 53
Db 208 NSFTEKYETEG--QRINKKVARERAKELPPKSSRPOPAH 249

```

RESULT 12

```

ID 090865 PRELIMINARY; PRT; 704 AA.
AC 090865;

```

```

DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR-LIKE/MACROPHAGE STIMULATING PROTEIN.

```

```

GN HGFI/MSF.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Bally S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGFI/MSF, and c-met suggests new functions
during early chick development.";
RL Dev. Genet. 17:90-101(1995).
RL EMBL: X84043; CAA5862.1; -.
DR HSSP: P00747; 1CEA.
DR MEROPS: S01.977; -.
DR INTERPRO: IPR000001; -.
DR INTERPRO: IPR001254; -.
DR INTERPRO: IPR001314; -.
DR INTERPRO: IPR003014; -.
DR PFAM: PF000024; PAN; 1.
DR PFAM: PF000051; kringler; 4.
DR PFAM: PF000089; trypsin; 1.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_4.
DR PROSITE: PS50070; KRINGLE_2; 4.
SQ SEQUENCE 704 AA: 79341 MW: CAB0D8CC41367C37 CRC64;

```

```

Query Match 20.5%; Score 64.5; DB 13; Length 704;
Best Local Similarity 23.8%; Pred. No. 6.1;
Matches 15; Conservative 7; Mismatches 24; Indels 17; Gaps 3;

```

```

OY 2 NPKHNC-----LQSCNSRSDYRNQACHRCMLKVEKECEEGELPRPRPOH 52
Db 173 NVRHQSGIKKCDVACMTNGE--DYRGFVDHTESG-----TECQRMWLDQPHKPHYH 224

```

```

OY 53 PER 55
Db 225 PDK 227

```

RESULT 13

```

ID 020155 PRELIMINARY; PRT; 414 AA.
AC 020155;

```

```

DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE F38B7.1 PROTEIN.
GN F38B7.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Lemard N.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -I- ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) MAY BE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE. THE SEQUENCE SHOWN IS THAT
CC OF FORM A.
DR EMBL: Z74033; CAA98475.1; -.
DR EMBL: Z74033; CAA98476.1; -.
DR INTERPRO: IPR000571; -.
DR PFAM: PF00642; zf-CCH; 2.
DR Hypothetical protein; Alternative splicing.
FT VARSPLIC 1 38 MISSING (IN ISOFORM B).
FT VARSPLIC 39 52 MISSING (IN ISOFORM B).
SQ SEQUENCE 414 AA: 44308 MW: D8B29D29A9E39583 CRC64;

```

```

Query Match          20.2%; Score 63.5; DB 5; Length 414;
Best Local Similarity 35.7%; Pred. No. 5.1;
Matches 20; Conservative 7; Mismatches 14; Indels 15; Gaps 5;

OY 1 ENPKHNCLOSNSERDSYRNQACHARCNLKVEKEBCE--ESEEIPRPPRPOHPE 54
      ||| | : | | | | | | | | | | | | | | | | | | | | | | | |
DB 194 KNPBYK-TELCSRWMD-----HGRCNY----GRCQYAHGELER-RVPRHPK 236

RESULT 14
O62446 PRELIMINARY: PRT; 663 AA.
ID 062446
AC 062446;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE Y43F4A.1 PROTEIN.
GN Y43F4A.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Ritken L., Roop A., Saunders D., Showkhen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL: Z99271; CAB16471.1; -.
DR INTERPRO: IPR000130; -.
DR INTERPRO: IPR001577; -.
DR PFAM: PF01457; Peptidase_M8; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 663 AA; 76049 MW; 814C31728B5ACEA CRC64;

Query Match          20.0%; Score 63; DB 5; Length 663;
Best Local Similarity 35.9%; Pred. No. 8.8;
Matches 14; Conservative 4; Mismatches 17; Indels 4; Gaps 1;

OY 7 KCLQSCNSERDSYRNQACHARCNLKVEKECEGEIPR 45
      || | : | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 KCISFYKYGKGRNVACDKGR---ERTTCGEADIPR 155

RESULT 15
O9KK20 PRELIMINARY: PRT; 519 AA.
ID O9KK20;
AC O9KK20;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE SURFACE PROTEIN PSpC.
GN PSpC.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.

```

```

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=699;
RA Tannelli F., Oggioni M.R., Spinoza M.R., Pozzi G.;
RT "Allelic variants of surface protein PspC (SpSA, CppA, PbcA) of
RT Streptococcus pneumoniae."
RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF154036; AAF73808.1; -.
SQ SEQUENCE 519 AA; 58731 MW; 7E1879295825AFD1E CRC64;

```

```

Query Match          19.8%; Score 62.5; DB 2; Length 519;
Best Local Similarity 27.5%; Pred. No. 8.1;
Matches 19; Conservative 11; Mismatches 20; Indels 19; Gaps 3;

OY 3 PHNKCLQSCNSERDSYRNQACHARCNLKVEKEBCE-----GEIPR 46
      | : : : | | | | | | | | | | | | | | | | | | | | | |
DB 211 PRDEQIKQAEAVES--KQAEATRLKIKITDREAEAEKRRADAKEDGKPRPAEQPQ 268
OY 47 RPRPOHPE 55
      | | | | |
DB 269 APAPO-PEK 276

```

Search completed: March 1, 2001, 16:17:18
Job time: 456 sec